

WHAT IS CLAIMED IS:

1. A speech recognition system comprising:
input means for inputting acoustic information;
analysis means for analyzing the acoustic
5 information input by said input means to acquire feature
quantity parameters;
first holding means for obtaining and holding
processing information for encoding on the basis of the
feature quantity parameters obtained by said analysis
10 means;
second holding means for holding processing
information for a speech recognition process in accordance
with the processing information for encoding;
conversion means for compression-encoding the
15 feature quantity parameters obtained via said input means
and said analysis means on the basis of the processing
information for encoding; and
recognition means for executing speech recognition
on the basis of the processing information for speech
20 recognition held by said holding means, and the feature
quantity parameters compression-encoded by said conversion
means.
2. The system according to claim 1, wherein said system
is built by a first apparatus having said analysis means,
25 said first holding means, and said conversion means, and
a second apparatus having said recognition means, and

said system further comprises communication means for sending the processing information generated by said first holding means and data acquired by said conversion means from the first apparatus to the second apparatus.

- 5 3. The system according to claim 1, wherein said second holding means holds processing information used to decode information converted by said conversion means, and

said recognition means comprises:

- 10 decoding means for decoding the compression-encoded feature quantity parameters by looking up the processing information held in said second holding means, and

said recognition means executes a speech recognition process on the basis of the feature quantity parameters decoded by said decoding means.

- 15 4. The system according to claim 2, wherein said second holding means is arranged in the second apparatus.

5. The system according to claim 1, wherein said second holding means makes some steps of a likelihood calculation associated with speech recognition on the basis of the
20 processing information for encoding and an acoustic model, and holds the calculation result as the information for speech recognition, and

- said recognition means obtains a speech recognition result by making a likelihood calculation for data acquired
25 by said conversion means using the information held by said second holding means.

6. The system according to claim 1, further comprising mode designation means for selectively executing a learning mode of making said first and second holding means function, and a speech recognition mode of making said conversion means and said recognition means function.

7. The system according to claim 1, wherein said conversion means scalar-quantizes multi-dimensional speech parameters obtained by said analysis means for respective dimensions.

8. The system according to claim 7, wherein the scalar quantization uses an LBG algorithm.

9. The system according to claim 7, wherein the scalar quantization assumes that data to be quantized form a Gaussian distribution, and quantizes with quantization steps having equal probabilities in the distribution.

10. The system according to claim 7, wherein setting means changes clustering for the scalar quantization on the basis of the feature quantity parameters obtained by said analysis means.

11. A speech recognition method comprising:
the input step of inputting acoustic information;
the analysis step of analyzing the acoustic information input in the input step to acquire feature quantity parameters;

the first holding step of obtaining processing information for encoding on the basis of the feature

quantity parameters obtained in the analysis step, and storing the information in first storage means;

the second holding step of holding, in second storage means, processing information for a speech recognition
5 process in accordance with the processing information for encoding;

the conversion step of compression-encoding the feature quantity parameters obtained via the input step and the analysis step on the basis of the processing information
10 for encoding; and

the recognition step of executing speech recognition on the basis of the processing information for speech recognition held in said second storage means in the second holding step, and the feature quantity parameters
15 compression-encoded in the conversion step.

12. The method according to claim 11, wherein a system is built by a first apparatus which executes the analysis step, the first holding step, and the conversion step, and a second apparatus which executes the recognition step, and
20 said method further comprises the communication step of sending the processing information generated in the first holding step and data acquired in the conversion step from the first apparatus to the second apparatus.

13. The method according to claim 11, wherein the second
25 holding step includes the step of holding, in said second storage means, processing information used to decode information converted in the conversion step, and

the recognition step comprises:

the decoding step of decoding the
compression-encoded feature quantity parameters by looking
up the processing information held in said second storage
5 means, and

the recognition step includes the step of executing
a speech recognition process on the basis of the feature
quantity parameters decoded in the decoding step.

14. The method according to claim 12, wherein the second
10 holding step is executed by the second apparatus.

15. The method according to claim 11, wherein the second
holding step includes the step of making some steps of a
likelihood calculation associated with speech recognition
on the basis of the processing information for encoding and
15 an acoustic model, and holding the calculation result as
the information for speech recognition, and

the recognition step includes the step of obtaining
a speech recognition result by making a likelihood
calculation for data acquired in the conversion step using
20 the information held in the second holding step.

16. The method according to claim 11, further comprising
the mode designation step of selectively executing a
learning mode of making the first and second holding steps
function, and the speech recognition mode of making the
25 conversion step and the recognition step function.

17. The method according to claim 11, wherein the
conversion step includes the step of scalar-quantizing

multi-dimensional speech parameters obtained in the analysis step for respective dimensions.

18. The method according to claim 17, wherein the scalar quantization uses an LBG algorithm.

5 19. The method according to claim 17, wherein the scalar quantization assumes that data to be quantized form a Gaussian distribution, and quantizes with quantization steps having equal probabilities in the distribution.

20. The method according to claim 17, wherein the setting
10 step includes the step of changing clustering for the scalar quantization on the basis of the feature quantity parameters obtained by the analysis step.

21. An information processing apparatus comprising:

input means for inputting acoustic information;
15 analysis means for analyzing the acoustic information input by said input means to acquire feature quantity parameters;

holding means for generating and holding processing information for compression-encoding on the basis of the
20 feature quantity parameters obtained by said analysis means;

first communication means for sending the processing information generated by said holding means to an external apparatus;

25 conversion means for compression-encoding the feature quantity parameters of the acoustic information

obtained via said input means and said analysis means on the basis of the processing information; and

second communication means for sending data obtained by said conversion means to the external apparatus.

5 22. An information processing apparatus comprising:

first reception means for receiving processing information associated with compression-encoding from an external apparatus;

holding means for holding, in a memory, processing
10 information for speech recognition obtained on the basis of the processing information received by said first reception means;

second reception means for receiving compression-encoded data from the external apparatus; and

15 recognition means for executing speech recognition of the data received by said second reception means using the processing information held in said holding means.

23. The apparatus according to claim 21, wherein said recognition means comprises:

20 decoding means for decoding data received by said second reception means using the processing information held in said holding means; and

means for executing a speech recognition process on the basis of feature quantity data decoded by said decoding
25 means.

24. The apparatus according to claim 21, wherein said holding means generates likelihood information on the basis

of the processing information received by said first reception means, and a predetermined acoustic model, and holds the likelihood information in the memory, and

said recognition means makes speech recognition by
5 making a likelihood calculation on the basis of data received by said second reception means using the likelihood information held in the memory.

25. An information processing method comprising:

the input step of inputting acoustic information;
10 the analysis step of analyzing the acoustic information input in the input step to acquire feature quantity parameters;

the holding step of generating and holding processing information for compression-encoding on the basis of the
15 feature quantity parameters obtained in the analysis step;

the first communication step of sending the processing information generated in the holding step to an external apparatus;

the conversion step of compression-encoding the
20 feature quantity parameters of the acoustic information obtained via the input step and the analysis step on the basis of the processing information; and

the second communication step of sending data obtained in the conversion step to the external apparatus.

25 26. An information processing method comprising:

the first reception step of receiving processing information associated with compression-encoding from an external method;

the holding step of holding, in a memory, processing
5 information for speech recognition obtained on the basis of the processing information received in the first reception step;

the second reception step of receiving
compression-encoded data from the external method; and

10 the recognition step of executing speech recognition of the data received in the second reception step using the processing information held in the holding step.

27. The method according to claim 26, wherein the recognition step comprises:

15 the decoding step of decoding data received in the second reception step using the processing information held in the holding step; and

the step of executing a speech recognition process on the basis of feature quantity data decoded in the
20 decoding step.

28. The method according to claim 26, wherein the holding step includes the step of generating likelihood information on the basis of the processing information received in the first reception step, and a predetermined acoustic model,
25 and holding the likelihood information in the memory, and

the recognition step includes the step of making speech recognition by making a likelihood calculation on

the basis of data received in the second reception step using the likelihood information held in the memory.

29. A computer readable medium for storing a control program for making a computer execute a speech recognition
5 process, said speech recognition process comprising:

the input step of inputting acoustic information;

the analysis step of analyzing the acoustic
information input in the input step to acquire feature
quantity parameters;

10 the first holding step of obtaining processing
information for encoding on the basis of the feature
quantity parameters obtained in the analysis step, and
storing the information in first storage means;

the second holding step of holding, in second storage
15 means, processing information for a speech recognition
process in accordance with the processing information for
encoding;

the conversion step of compression-encoding the
feature quantity parameters obtained via the input step and
20 the analysis step on the basis of the processing information
for encoding; and

the recognition step of executing speech recognition
on the basis of the processing information for speech
recognition held in said second storage means in the holding
25 step, and the feature quantity parameters
compression-encoded in the conversion step.

30. A computer readable medium for storing a control program for making a computer execute a predetermined information process, said predetermined information process comprising:

5 the input step of inputting acoustic information;
 the analysis step of analyzing the acoustic information input in the input step to acquire feature quantity parameters;

 the holding step of generating and holding processing
10 information for compression-encoding on the basis of the feature quantity parameters obtained in the analysis step;

 the first communication step of sending the processing information generated in the holding step to an external apparatus;

15 the conversion step of compression-encoding the feature quantity parameters of the acoustic information obtained via the input step and the analysis step on the basis of the processing information; and

 the second communication step of sending data
20 obtained in the conversion step to the external method.

31. A computer readable medium for storing a control program for making a computer execute a speech recognition process, said speech recognition process comprising:

 the first reception step of receiving processing
25 information associated with compression-encoding from an external method;

the holding step of holding, in a memory, processing information for speech recognition obtained on the basis of the processing information received in the first reception step;

- 5 the second reception step of receiving compression-encoded data from the external apparatus; and
the recognition step of executing speech recognition of the data received in the second reception step using the processing information held in the holding step.